



# Une interface de visualisation avec retour de pertinence pour la recherche d'images

Thi Lan Le

## ► To cite this version:

Thi Lan Le. Une interface de visualisation avec retour de pertinence pour la recherche d'images. Rencontres des Jeunes Chercheurs en Recherche d'Information, Mar 2006, Lyon, France. inria-00186399

**HAL Id: inria-00186399**

**<https://inria.hal.science/inria-00186399>**

Submitted on 9 Nov 2007

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

---

# Interface de visualisation avec retour de pertinence pour la recherche d'images

Le Thi Lan<sup>\*,\*\*</sup>

<sup>\*</sup> *Projet ORION, INRIA, 2004 route des Lucioles – B.P. 93  
06902 Sophia Antipolis Cedex, FRANCE  
Lan.Le\_Thi@sophia.inria.fr*

<sup>\*\*</sup> *Centre MICA, Institut Polytechnique de Hanoi, VIETNAM*

---

**RÉSUMÉ.** *Le domaine de la recherche d'images est encore un problème ouvert en comparaison avec les résultats obtenus dans le domaine de la recherche de textes. C'est encore plus vrai en ce qui concerne la recherche d'images par le contenu sémantique. Dans cet article, nous proposons une nouvelle technique de recherche d'images permettant, par la communication avec l'utilisateur, d'extraire de façon rapide et conviviale son objectif. Ici nous proposons une interface graphique 2D adaptée au problème de la recherche d'images permettant une communication bidirectionnelle : du système vers l'utilisateur pour lui visualiser les résultats courants de la recherche et de l'utilisateur vers le système pour qu'il fournisse des informations de type retour de pertinence pour raffiner sa requête. Le retour de pertinence est uniquement sous forme de validation d'images résultats (ou exemples positifs). Cette approche a été implémentée et testée sur une base de données de 1000 images. Les résultats préliminaires sont prometteurs. Une comparaison avec une interface traditionnelle est montrée.*

**ABSTRACT.** *The domain of image retrieval is still an open problem in comparison with the results obtained in the domain of text retrieval. The problem of semantic image retrieval is even more difficult and open. In this article, we propose a new technique of image retrieval enabling, by interaction with the user, to extract fast and easily his/her objective. Here we propose a 2D graphic interface adapted to the problem of image retrieval and enabling a bidirectional communication: from the system towards the user to visualize the current research results and from the user towards the system so that the user can provide some relevance feedback information to refine his/her query. Relevance feedback is only in the form of validation of result images (or positive examples). This approach has been implemented and tested on an image database of 1000 images. The results are promising. A comparison with a traditional interface is shown.*

**MOTS-CLÉS :** *Indexation et Recherche d'images par le contenu, Visualisation, Retour de pertinence.*

**KEYWORDS:** *Content-Based Image Indexing and Retrieval, Visualization, Relevance Feedback*

---

## 1. Introduction

La recherche d'images par le contenu semble dater de 1992, quand elle a été employée par T. Kato [KAT 92]. Quelques résultats ont été obtenus depuis mais c'est encore un problème ouvert en comparaison avec les résultats obtenus dans le domaine de la recherche de textes. C'est encore plus vrai en ce qui concerne la recherche d'images par le contenu sémantique. En effet, les bases de données d'images sont souvent indexées par des attributs bas niveaux et non par leur contenu sémantique. Il existe alors un fossé sémantique, à savoir une différence importante, entre la similarité calculée à partir des index et la similarité attendue par l'utilisateur [SME 00]. Pour combler ce fossé sémantique dans ce domaine, deux approches sont possibles : soit indexer les images par le contenu sémantique, soit améliorer la technique de recherche d'images par des échanges avec l'utilisateur. Nous nous intéressons, ici à la deuxième approche. Pour améliorer les échanges avec l'utilisateur, deux défis doivent être relevés :

- L'amélioration du noyau du système (méthode de retour de pertinence) pour mieux correspondre aux attentes de l'utilisateur.
- La visualisation qui permet au système de présenter des résultats de recherche à l'utilisateur et de lui demander des informations.

Concernant le premier défi, beaucoup de travaux ont été réalisés. Ils proposent soit de modifier la requête et la distance utilisée dans le calcul de similarité [RUI 98] soit de tenter d'approximer les surfaces de décision séparant des images pertinentes et des images non-pertinentes dans l'espace des attributs [ZHA 05]. Concernant le deuxième défi, l'interface traditionnelle (l'affichage de l'image requête et d'un nombre fixé d'images résultats sous la forme d'une liste d'images) devient inappropriée dans un système interactif de recherche d'images. Depuis ces dernières années, la visualisation attire beaucoup l'attention des chercheurs dans le domaine de la recherche d'images et quelques méthodes de visualisation ont été proposées. Dans [RUB 98], l'utilisation de la distance EMD (Earth Mover's Distance) et de MDS (Multidimensional Scaling) donne une possibilité de présenter des images résultats sur un plan 2D selon une seule mesure de similarité. Il manque la capacité de faire le retour de pertinence. En 2004, Deng [DEN 04] a proposé une méthode de visualisation et de comparaison des images basée sur les cartes auto-organisatrices. Au début, cette méthode ne fournit pas de retour de pertinence. À l'heure actuelle, quelques travaux en cours [LAA 05] tentent de lui ajouter un retour de pertinence.

Nos travaux visent à résoudre ces deux défis en même temps par la conception d'une interface et d'une méthode de retour de pertinence correspondant à cette interface.

Dans la section 2 de cet article, nous présentons notre interface graphique 2D et notre méthode de retour de pertinence correspondant. La section 3 est consacrée aux expérimentations et aux comparaisons. La section 4 conclut cet article par les perspectives envisagées pour perfectionner cette technique.

## 2. Interface graphique 2D avec retour de pertinence

### Interface graphique 2D proposée

Notre idée est de concevoir une interface qui permet de rassembler autour de l'image requête des images qui lui sont semblables. L'utilisateur sélectionne alors une région sur l'écran pour raffiner sa requête. La figure 2.a montre notre interface graphique 2D. Cette interface se divise en 5 zones : zone de fonctions, zone d'exemples initiaux, zone d'exemples négatifs, zone d'exemples positifs et zone de résultats. La zone de résultats est la zone la plus importante. Dans cette zone, l'image requête R est au milieu, les images résultats sont affichées selon deux axes : l'un utilisant l'histogramme couleur dans les espaces Rouge Vert Bleu (RVB) et l'autre utilisant celui dans les espaces Teinte Saturation Valeur (TSV). La position (x,y) d'une image I dans la base d'images est calculée comme suit :

$$\begin{aligned}x(I) &= d_{Intersection}(H_{RVB}(I), H_{RVB}(R)) \\y(I) &= d_{Intersection}(H_{TSV}(I), H_{TSV}(R)) \\sign(x) &= 1 \text{ si } E(H_{RVB}(I)) < E(H_{RVB}(R)) \text{ et } sign(x) = -1 \text{ inversement} \\sign(y) &= 1 \text{ si } E(H_{TSV}(I)) < E(H_{TSV}(R)) \text{ et } sign(y) = -1 \text{ inversement} \quad [1]\end{aligned}$$

où : E est l'entropie,  $H_{RVB}$ ,  $H_{TSV}$  sont les histogrammes dans l'espace de couleur RVB et TSV  $d_{Intersection}$  est la distance entre deux histogrammes [SWA 91]. Avec cette interface, à l'heure actuelle, nous n'employons que des attributs globaux. Des attributs locaux peuvent être utilisés pour améliorer les résultats.

### Retour de pertinence avec interface graphique 2D

La méthode de retour de pertinence présentée dans cette section consiste à faire le retour de pertinence par modification de la requête. Cette méthode marche uniquement avec des exemples positifs (validation). Dès que des résultats courants sont affichés sur l'écran, l'utilisateur peut cliquer sur une image pour la sélectionner comme exemple positif, et cette image va s'ajouter dans la zone des exemples positifs qui est initialisé par la requête initiale. La nouvelle requête R' est un ensemble de n exemples positifs  $P_j$ . La nouvelle position (x',y') de l'image I est calculée comme :

$$\begin{aligned}x'(I) &= d(I, R') = \min(d_{Intersection}(I, P_j) \text{ avec } j \in (1, n)) \\y'(I) &= d(I, R') = \min(d_{Intersection}(I, P_j) \text{ avec } j \in (1, n)) \quad [2]\end{aligned}$$

Nous garderons les valeurs de sign(x) et sign(y) pour sign(x') et sign(y'). Les résultats suivants seront affichés selon leurs nouvelles positions. L'utilisateur peut faire ce processus plusieurs fois jusqu'à ce qu'il soit satisfait.

## 3. Expérimentations

### Base d'images utilisée

Pour évaluer et valider notre système, nous avons utilisé un sous-ensemble (<http://wang.ist.psu.edu/docs/related/>) de la base d'images Corel. Ce sous-ensemble contient 1000 images naturelles en couleurs divisées en 10 classes. Chaque image

dans cette base d'images a une taille de  $384 \times 256$  pixels ou  $256 \times 384$  pixels.

### Expérimentations avec l'interface traditionnelle

Pour comparer les résultats avec notre interface graphique 2D, nous avons développé une interface traditionnelle (fig. 1.a). La méthode de retour de pertinence [RUI 98] a été appliquée. Dans cette interface, l'image requête est à gauche, les seize premières images résultats sont affichées à droite, sous chaque image résultat, on a trois options à choisir (pertinente, neutre, non-pertinente) pour faire le retour de pertinence. Avec la méthode de Rui, les deux descripteurs utilisés sont la couleur et la texture (quatre représentations utilisées pour la couleur : l'histogramme [SWA 91](RVB et TSV), les moments [STR 95](RVB et TSV) et une représentation utilisée pour la texture : les matrices de co-occurrences [HAR 73]). La figure 1 montre un résultat de cette méthode avec l'interface traditionnelle. Dans cet exemple, l'utilisateur veut chercher des images contenant un (des) cheval (chevaux). L'image (a) montre les seize premiers résultats, les images en rouge sont les images non pertinentes, les autres sont pertinentes. L'image (b) montre les seize premiers résultats quand on fait une fois le retour de pertinence. Le nombre d'images non-pertinentes diminue et leurs rangs sont beaucoup plus élevés (pertinence faible) que ceux des images pertinentes (premiers rangs).

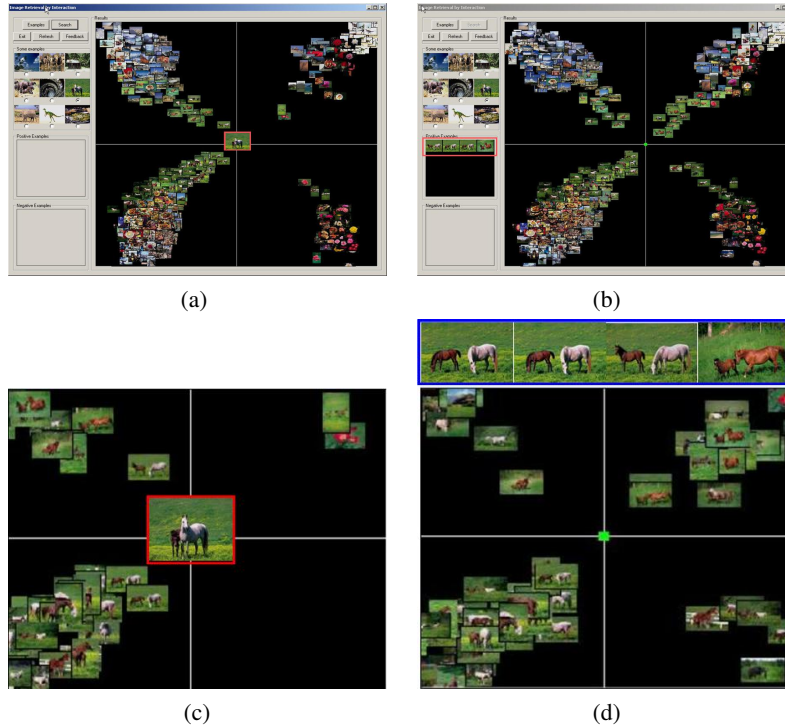


**Figure 1.** (a) Les seize premiers résultats sans retour de pertinence, les images en rouge sont les images non pertinentes, les autres sont pertinentes (b) Les seize premiers résultats quand on fait une fois le retour de pertinence

### Expérimentations avec l'interface graphique 2D

La figure 2 montre un résultat de notre interface avec le retour de pertinence correspondant. L'intention de l'utilisateur est de chercher des images contenant un (des) cheval (chevaux). La figure 2.a montre des résultats de recherche sans retour de pertinence, la figure 2.b montre de nouveaux résultats dans ce cas, l'utilisateur a choisi quatre exemples positifs. Les nouveaux résultats montrent une amélioration après avoir affectué un seul retour de pertinence. La plupart des images autour de l'image requête contiennent un (des) cheval (chevaux).

### Quelques comparaisons



**Figure 2.** (a) Des résultats sans retour de pertinence, l'image requête est au milieu, les images résultats sont affichées selon deux axes : l'axe vertical utilisant l'histogramme couleur RVB et l'axe horizontal utilisant l'histogramme couleur TSV (b) Des résultats après une fois le retour de pertinence, quatre exemples positifs sont choisis (c) Un zoom d'une région centrale des résultats sans retour de pertinence (d) Un zoom d'une région centrale des résultats après une fois de retour de pertinence

À partir des observations sur les résultats obtenus à partir de 2 interfaces, nous donnons ici quelques comparaisons :

- L'interface traditionnelle présente des résultats d'une façon uni-dimensionnelle alors que notre interface graphique 2D les présente d'une façon bi-dimensionnelle. Cela permet de montrer plus d'informations que via l'interface traditionnelle. L'interface graphique 2D donne implicitement la similarité entre plusieurs images résultats (l'interface traditionnelle ne donne que la similarité entre l'image résultat et l'image requête) ;
- L'interface graphique permet de grouper des images semblables (ex : dans la figure 2.a le groupe de fleurs est en bas à droite) ;
- L'interface traditionnelle limite le nombre de résultats affichés, elle rend difficile à l'utilisateur de fournir suffisamment d'exemples positifs et négatifs. Notre interface

graphique est capable de résoudre ce problème.

#### 4. Conclusions et perspectives

Dans cet article, nous avons présenté une interface graphique 2D et une méthode de retour de pertinence. Les comparaisons montrées dans la section précédente prouvent l'efficacité de notre technique. Cependant, il y a encore quelques problèmes à résoudre pour améliorer cette technique, tels que : l'affichage des images lorsque la base d'images devient très grande ; le processus de retour de pertinence pour cette interface avec des exemples négatifs. Nous avons fait cette interface en 2D, cela signifie que nous utilisons 2 attributs de l'image. Cependant, nous avons d'autres attributs. Nous regardons dans deux directions, l'une est de manipuler un espace de dimension supérieure à 2, l'autre est de laisser l'interface 2D en utilisant la technique d'Analyse en Composantes Principales pour réduire le nombre de dimensions à afficher. Nos perspectives sont de résoudre ces problèmes.

#### 5. Bibliographie

- [DEN 04] DENG D. ZHANG J. P. M., « Visualisation and Comparison of Image Collections based on Self-organised Maps », *Proc. Information security, Data Mining and Web Intelligence, and Software Internationalisation*, 2004.
- [HAR 73] HARALICK R.M. SHANMUGAM K. D. I., « Textural Features for Image Classification », *IEEE Trans. On Systems, Man, and Cybernetics*, vol. SMC-3, 1973, p. 610-621.
- [KAT 92] KATO T., HIRATA K., « Query by visual example in content-based image retrieval », *Proc. EDB192. Lecture Notes in computer Science*, 1992, p. 56-71.
- [LAA 05] LAAKSONEN J., VIITANIEMI V., KOSKELA M., « Emergence of semantic concepts in visual databases », *In Proceedings of International and Interdisciplinary Conference on Adaptive Knowledge Representation and Reasoning (AKRR 05)*, 2005, p. 127-134.
- [RUB 98] RUBNER Y., TOMASI C., GUIBAS L. J., « A Metric for Distributions with Applications to Image Databases », *Proceedings of the Sixth International Conference on Computer Vision table of contents*, 1998.
- [RUI 98] RUI Y., HUANG T. S., ORTEGA M., MEHROTRA S., « Relevance Feedback : A Power Tool for Interactive Content-Based Image Retrieval », *IEEE Transaction On Circuits and Video Technology*, 1998.
- [SME 00] SMEULDERS A. W., WORRING M., SANTINI S., GUPTA A., JAIN R., « Content-Based Image Retrieval at the End of the Early Years », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, 2000, p. 1349-1380.
- [STR 95] STRICKER M., ORENGO M., « Similarity of Color Images », *In Proceedings of SPIE, Vol. 2420(Storage and Retrieval of Image and Video Databases III)*, 1995.
- [SWA 91] SWAIN M., BALLARD D., « Color indexings », *International Journal of Computer Vision*, vol. 7, n° 1, 1991, p. 11-22.
- [ZHA 05] ZHANG C., CHEN X., « Region-Based Image Clustering and Retrieval Using Multiple Instance Learning », *CIVR*, 2005, p. 194-204.